I&MP for Transport Infrastructure Management using Deep Reinforcement Learning

Shreya Kejriwal

Supervisors: Prof. C. Andriotis, Prof. M. Overend, P. Bhustali

TU Delft

Agenda

01
Introduction

02
Research Overview

03
Environment modelling

04
DRL and scenario setup

05
Results and Discussions

TU Delft

Transport network

Vital part of the urban systems that support

- Mobility of users and businesses
- Transfer of goods and services.

Impact of poor transport network management

- Decreased performance: high travel times
- Safety concerns
- Roadblock to sustainability

Network pressure

Investment Backlog [1]

Traffic Overload [2]

Ripple effect of Losses [3]

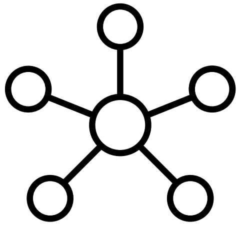
[1] https://www.nytimes.com/2020/12/22/world/europe/genoa-bridge-collapse.html (Retrieved on 15-05-2024)
[2] https://www.newcivilengineer.com/latest/significant-overload-caused-norways-timber-bridge-collapse-05-12-2022/ (Retrieved on 21-01-2024)
[3] https://www.news5cleveland.com/philadelphia-bridge-collapse-likely-to-have-broader-economic-impacts (Retrieved on 27-01-2024)

Inspection & Maintenance Planning

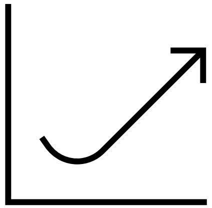
- Traditional Maintenance strategies fall short in
  - Identifying and predicting true health
  - Optimising timely repairs
  - Minimising disruptions

- Predictive maintenance
  - Informed data driven decisions
  - Prevent costly repairs and extend roadways lifespan

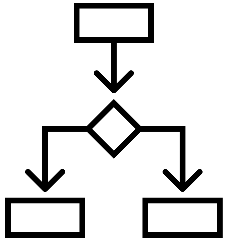
Research Focus



Individual component management to Network management



Consideration for fast changing environments



Consideration for multiple stakeholders and their preference

TU Delft

Agenda

01

Introduction

02

Research Overview

03

Environment modelling

04

DRL and scenario setup

05

Results and Discussions

TU Delft

Research Question

How can a **multi-asset transportation network** be effectively modeled to account for **dynamic influences** such as traffic variations, policy changes, and system interdependencies to **optimize network management?**

- How can we model the **multiple types** of transport network assets effectively?
- What's the best way to simulate a **changing traffic pattern**?
- What methods can be used to incorporate **stakeholder goals** into the reinforcement learning model?
- How effective are **DRL Methods** for developing I&M policies to manage the transportation network?
- What is the impact on different objectives (agencies, users, and societal factors/costs) and policies?

Methodology

Represent the transport system as a Graph

Define the network based on a Decision-making framework

Create a Reinforcement Learning Agent

Test on a developed Network Scenario

Analyse results, validate the returned Policy

Agenda

01

Introduction

02

Research Overview

03

Environment modelling

04

DRL and scenario setup

05

Results and Discussions

TU Delft

Graph Representation

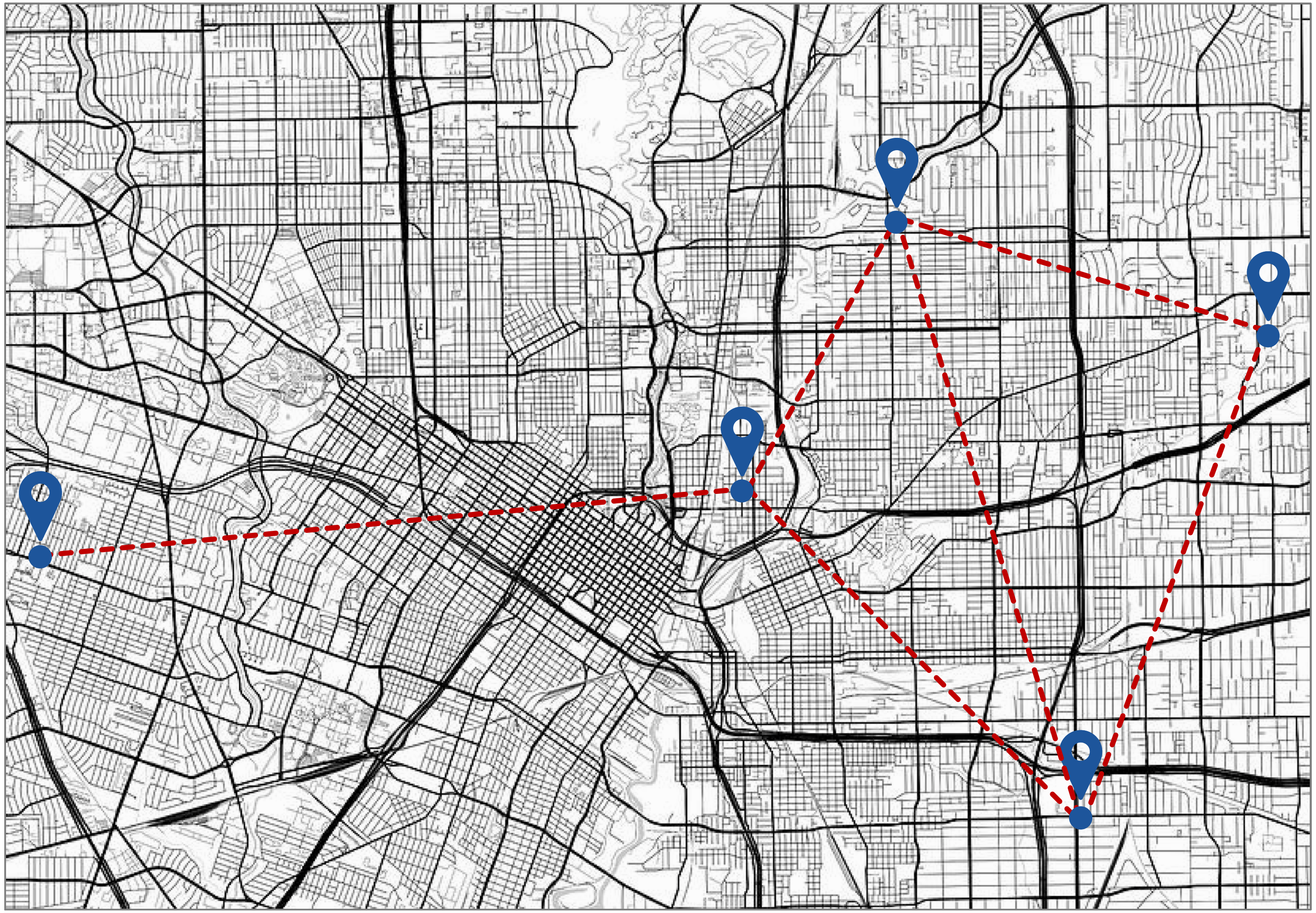


Fig 1: Representation of a transport network as a graph

Map by Jurq Studio

Graph Representation

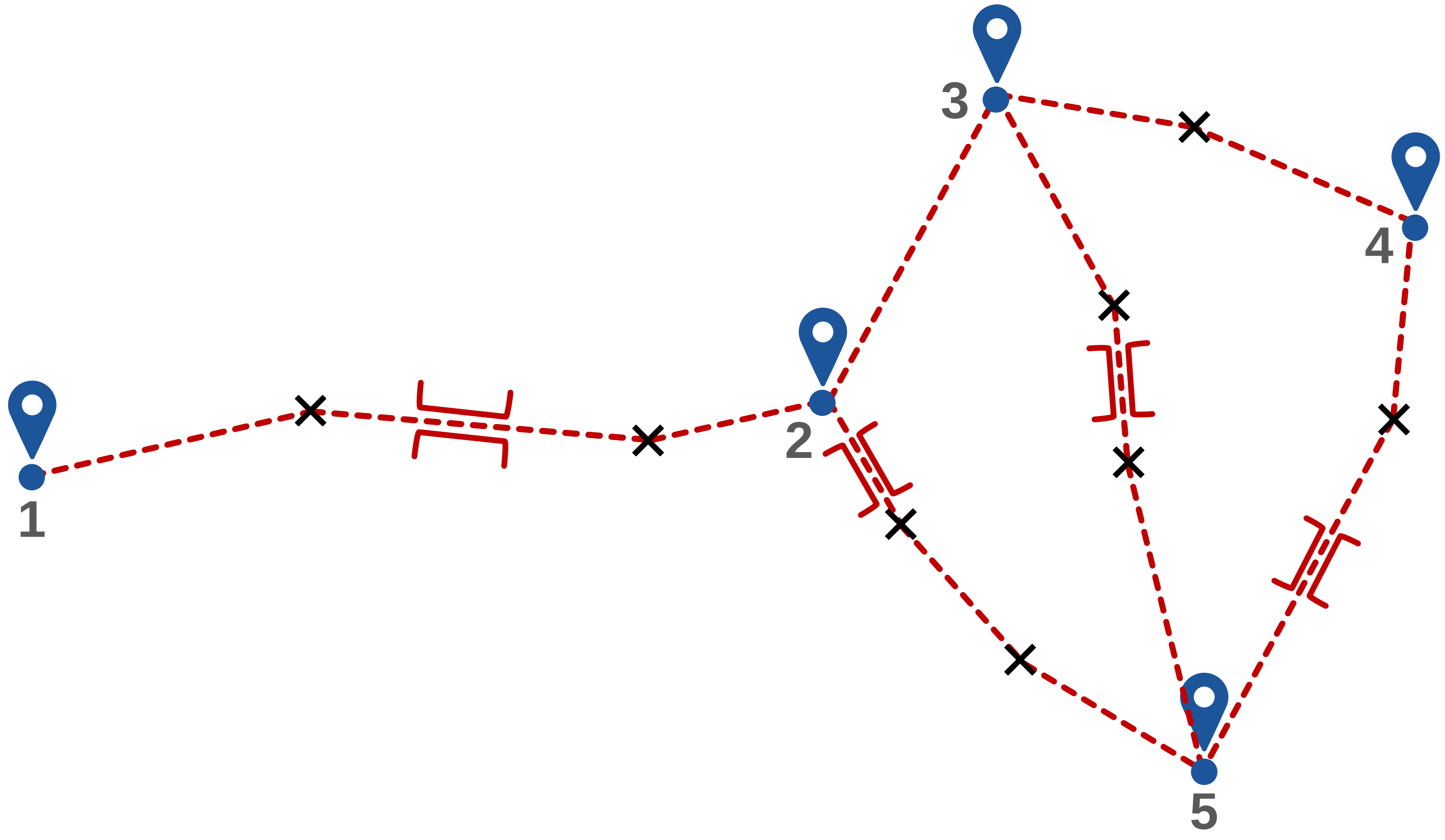
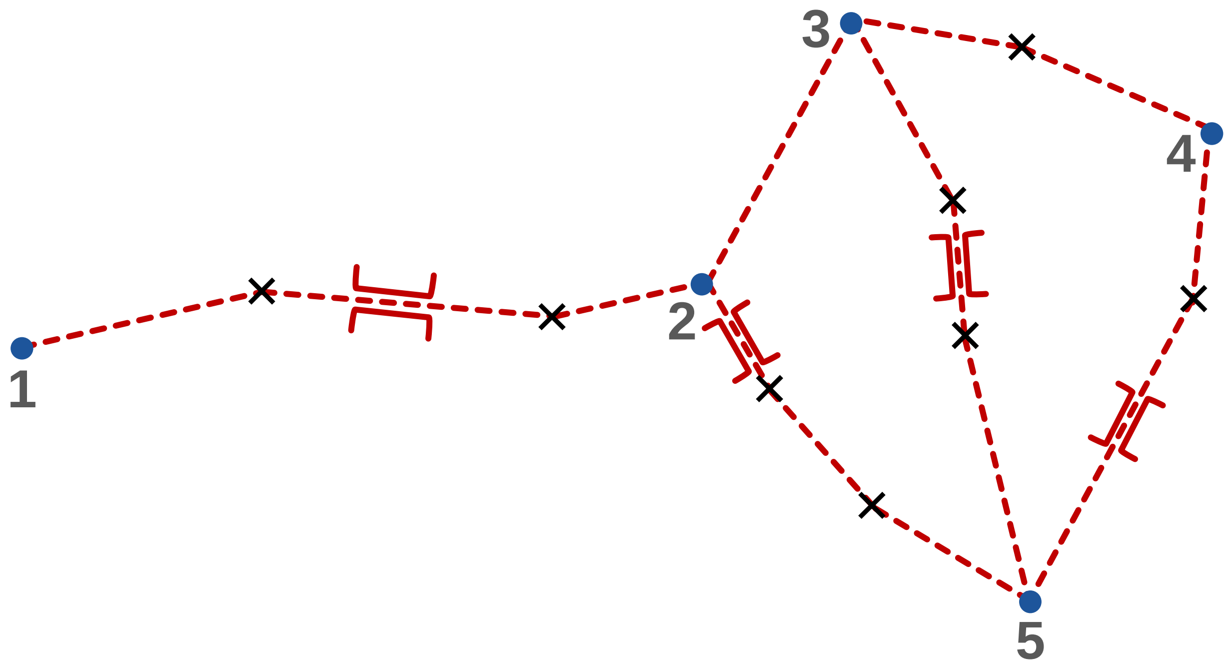


Fig 1: Representation of a transport network as a graph

Benefits of graph structure

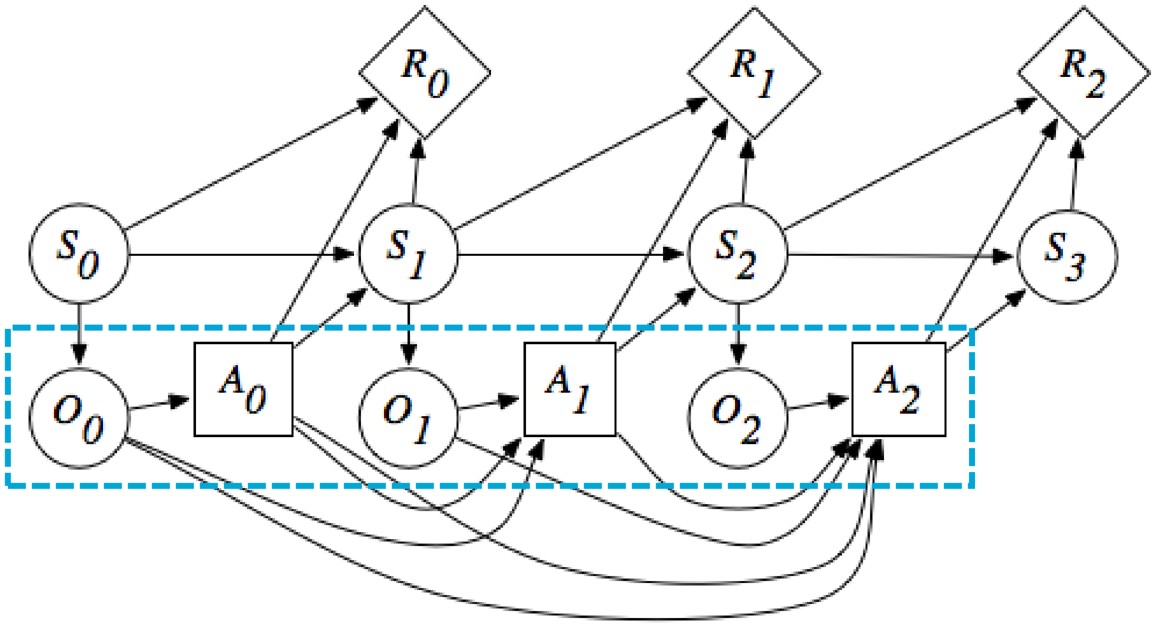
- Hierarchical structure:
  - Ease of adding and removing segments.
  - Considers different typologies, topologies and contexts.

- Precise maintenance actions and performance analysis.



Modelling the Network

- Markov Decision Process (MDP)
- Partially Observable Markov Decision Process (POMDP)
  - Accounting for uncertainties.
  - Uses beliefs and observations in the decision-making process.

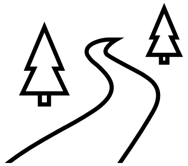


Poole, D. L., & Mackworth, A. K. (2023). Artificial Intelligence: Foundations of Computational Agents (3rd ed.). Cambridge University press.

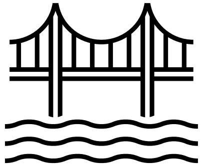
TU Delft

POMDP | Components

POMDP Tuple $< (P, B), A, T, \Omega, O, C, \gamma >$



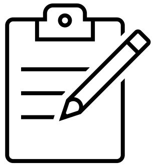
Pavement States (P)
Possible states of the pavement



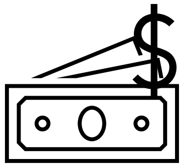
Bridge States (B)
Possible states of the bridges

Actions (A)
Set of possible actions for the agent

Transition Prob (T)
Probability of transitioning: $P(s'|s, a)$



Observation Prob (O)
Probability of observing; $O: S \ x \ A \ \rightarrow \ \pi(\Omega)$



Cost (C)
Cost accrued for taking an action in the state

Pavement States (P)

- 2 Features:
  - International Roughness Index (IRI) for surface roughness - Ride quality for users.
  - Critical Condition Index (CCI) - comprehensive assessment including structural condition
- Pavement is modelled as a single array of both beliefs [IRI, CCI]

State	Pavement condition	IRI (m/km)
$s_p = 5$	Very good	< 0.95
$s_p = 4$	Good	0.95 – 1.56
$s_p = 3$	Fair	1.57 – 2.19
$s_p = 2$	Mediocre	2.20 – 3.14
$s_p = 1$	Poor	> 3.15

State	Pavement condition	CCI
$s_p = 6$	Excellent	100 – 90
$s_p = 5$	Very good	89 – 80
$s_p = 4$	Good	79 – 61
$s_p = 3$	Fair	60 – 50
$s_p = 2$	Poor	49 - 37
$s_p = 1$	Very poor	< 37

Virginia Department of Transportation (2018). State of the Pavement. Retrieved from vdot.virginia.gov/media/vdotvirginiagov/about/highways

Bridge States (B)

- Only decks are considered:
  - In direct contact with the traffic.
  - Vulnerable to several weather-related deterioration and mechanical wear. [2]
- Modelled as an array
  - Observation padding

State	Bridge condition	Description[1]
$s_b = 9$	Excellent	NA
$s_b = 8$	Very good	No problems noted
$s_b = 7$	Good	Some minor problems noted
$s_b = 6$	Satisfactory	Minor section loss, cracking, spalling or scour
$s_b = 5$	Fair	Advanced section loss, deterioration, spalling or scour
$s_b = 4, \ldots$	Poor – Critical	Loss of section, deterioration, affects primary structural components, local failures and shear cracks
$s_b = 0$	Failed	Out of service

[1] FHWA (2004). Bridge conditions and performance. Retrieved from https://www.fhwa.dot.gov/policy/2004cpr/chap3c.cfm

[2] Saifullah et al., 2024

Action Space (A)

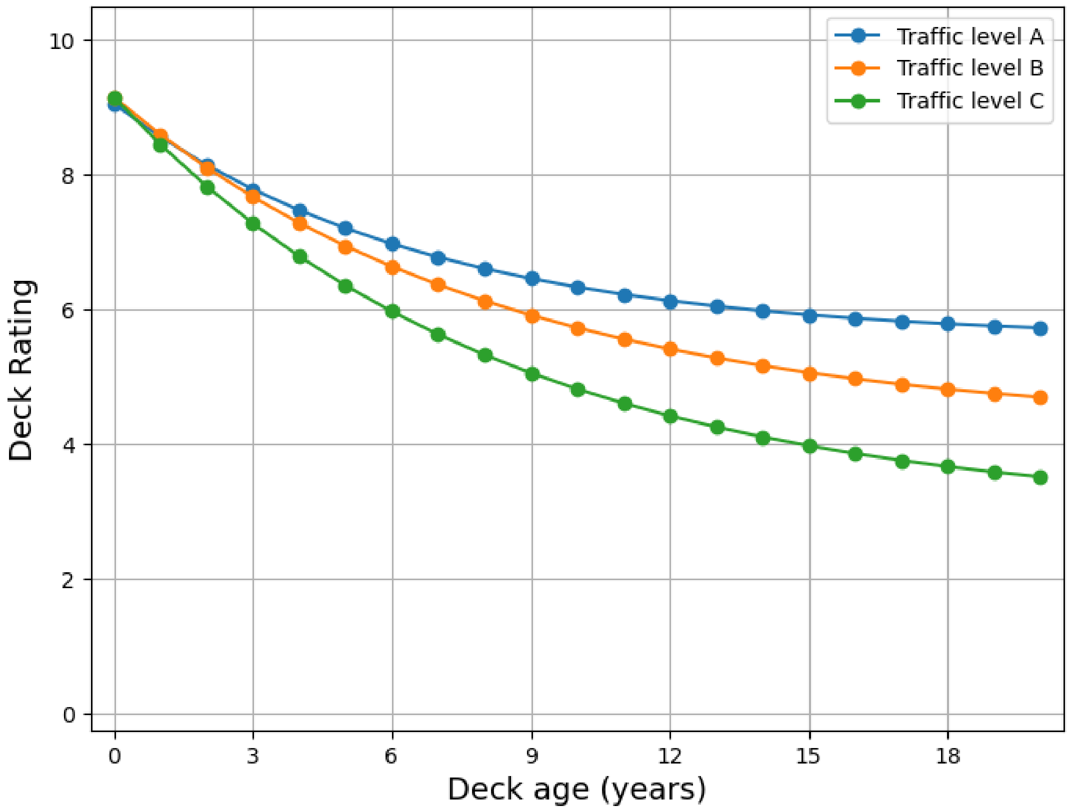
Action	Action name	Impact on Pavement / Deck
i_0	No inspection	Belief remains ambiguous. Observation error is ∞.
i_1	Routine inspection	Low fidelity inspection. Observation error is higher.
i_2	In-depth inspection	High fidelity inspection. Observation error is lower.
a_0	Do nothing	No action is performed. The segment continues to deteriorate.
a_1	Minor repair	Small improvements in the condition.
a_2	Major repair	Majorly improves the segments condition.
a_3	Reconstruction	Entire segment is reconstructed to its original state.

Action space = 3 x 4

= 12 discrete actions

Transition Probabilities (T)

- Nonstationary transition:
  - Critical Condition Index
  - Deck rating
- Influence of traffic change and age.
- IRI index has a stationary transition model.



Saifullah, M., Papakonstantinou, K. G., Andriotis, C. P., & Stoffels, S. M. (2024). Multi-agent deep reinforcement learning with centralized training and decentralized execution for transportation infrastructure management

Lou, P., Nassif, H., Su, D., and Truban, P. (2016). Effect of Overweight Trucks on Bridge Deck Deterioration Based on Weigh-in-Motion Data

Dynamic Traffic Conditions

- 20-year forecast of annual growth in trips: 0.5% – 0.9%.[1]
- Need consideration for distribution of vehicles.
- Steep increase in network volume
  - Heavy vehicles growing at a faster rate

Nodes	0	1	2	3	4
0	0	**10**	**8**	0	**9**
1	0	0	**9.9**	**7.2**	**18**
2	0	0	0	**7.9**	**24.8**
3	0	0	0	0	**15**
4	0	0	0	0	0

Table 1: Trips per day between nodes (thousand vehicles/ day)

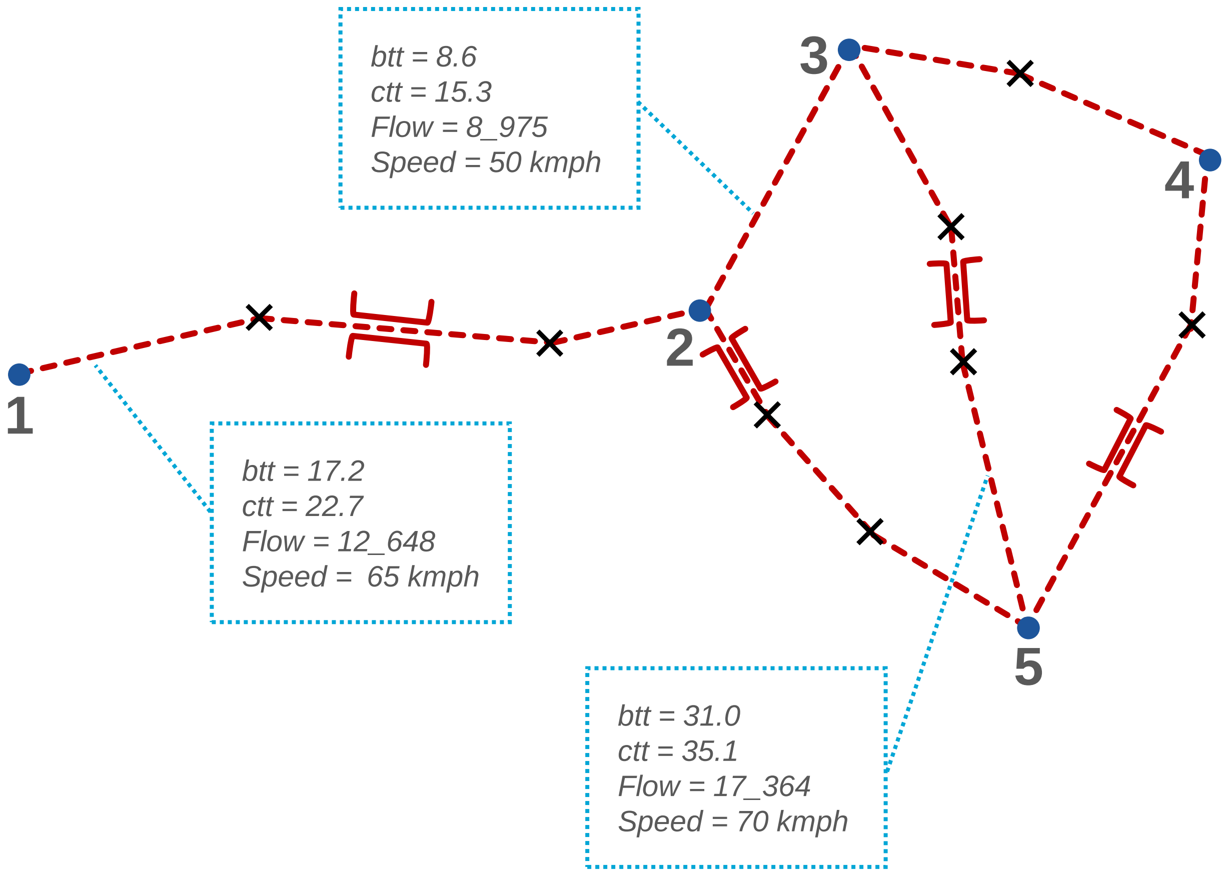
Vehicle	PCE	Distribution[1]	Growth%[2]
Motorcycle	0.50	0.28	0.50
Car	1.00	72.16	0.50
Bus	3.00	18.30	0.50
Truck	4.50	1.72	1.80
XL Truck	5.00	7.54	1.20

Table 2: Vehicles travelling in the network

[1] https://www.fhwa.dot.gov/policyinformation/statistics/2021/vm4.cfm3
[2] https://www.fhwa.dot.gov/policyinformation/tables/vmt/vmt_forecast

Traffic Model

- User Equilibrium method
- **Inputs:**
  - Graph, trips, capacity
- **Outputs at each timestep:**
  - Base Travel Time
  - New travel time
  - Flow of vehicles
  - Speed range per link
  - Level of Service of the network



Cost Model

Agency objective	*User objective*	*Env. objective*	*Safety objective*
- Department Of Transportation, - Maintenance agency	- Daily commuters, - Goods and service transport	- Climate department, - International organizations	- Department Of Transportation, - Asset Manager & - Engineers

TU Delft

Agency Cost

Economic implications of the maintenance and management:

a. Cost of maintaining and inspecting
b. Risk based cost due to probable failure
c. Additional maintenance cost due to overweight vehicles

This cost is calculated in $.

$$C_{agency}(t) = \frac{1}{(1+r)^t}\left(C_{direct}(t) + C_{indirect}(t) + C_{overweight}(t)\right)$$

User Cost

Financial burden of network condition and agency actions on users:

a. Delay cost due to planned maintenance actions
b. Vehicle operating cost

This cost is calculated in $.

$$C_{user}(t) = \frac{1}{(1+r)^t}\left(C_{delay}(t) + C_{veh\_operating}(t)\right)$$

Environmental Emissions

Carbon footprint generated due to management of the network directly and indirectly:

a. Maintenance actions
b. Detour
c. Congestion in the network

This is calculated in tones.

$$C_{env}(t) = C_{congestion}\,(t) + C_{action}\,(t) + C_{detour}(t)$$

Safety Rating

Measure the performance of the components as well as the entire network:

a. Safety condition of individual components
b. Failure probability of the entire network

$$C_{safety}(t) = (C_{comp}(t) + C_{network}(t))$$

Concerns about using many objectives

- Multiple scales and types of metrics
  - Cost, emissions, ratings.
- Difficult to combine and compare as a single cost function.
- Decision makers risk tendencies and preference

Multi-Attribute Utility Model

$$C(t) = w_{agency}.u_{agency}(t) + w_{user}.u_{user}(t) + w_{env}.u_{env}(t) + w_{safety}.u_{safety}(t)$$

- Defines utility between 0-1
- Weights for each metric to indicate preference or importance
- Crucial to define appropriate C_{metric}^{max} and $C_{metric}^{\min}$ values

Metric value	Utility function value
C_{metric} is close to C_{metric}^{min}	Utility value is close to 1
C_{metric} is close to C_{metric}^{max}	Utility value is close to 0
C_{metric} is more than C_{metric}^{max}	Utility value becomes negative

Ang, A.-S. and Tang, W. (1984). Probability Concepts in Engineering Planning and Design: Decision, Risk, and Reliability, volume II. Wiley Press.

Agenda

01
Introduction

02
Research Overview

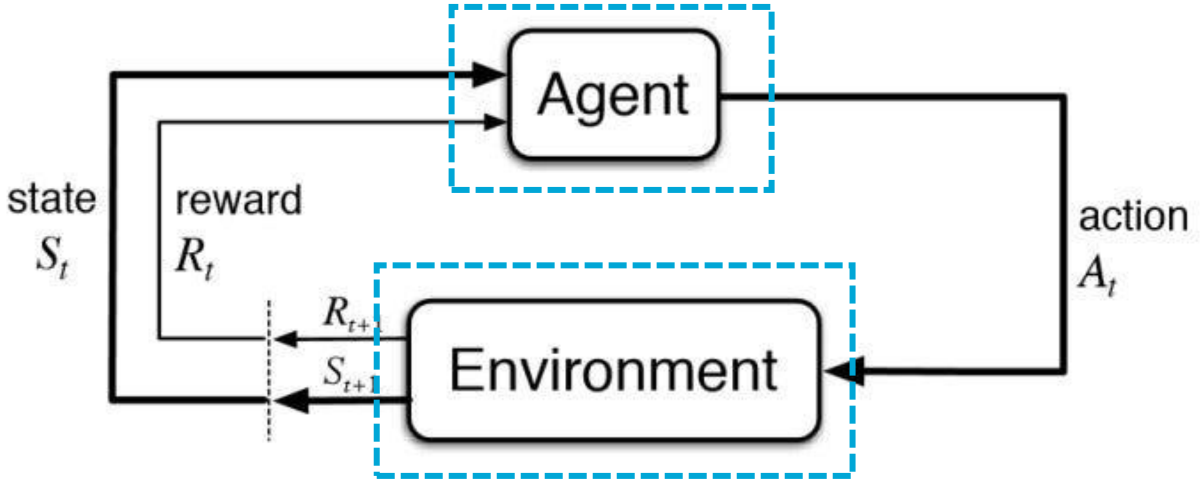
03
Environment modelling

04
DRL and scenario setup

05
Results and Discussions

Deep Reinforcement Learning

- Agent interacts with an environment to maximise cumulative rewards.
- Learns by trial and error.
- Balance between exploring new actions and exploiting existing knowledge



- Algorithms used: Double Deep Q-Learning Algorithm (DDQN) and Deep Centralised Multi-agent Actor Critic (DCMAC)
- Online implementation from: IMPRL Github

TU Delft

Deep Reinforcement Learning

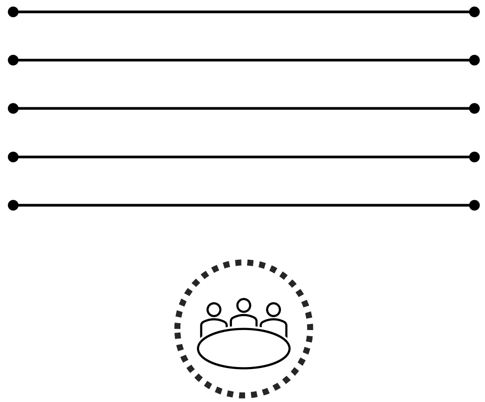
Double Deep Q-Learning Algorithm

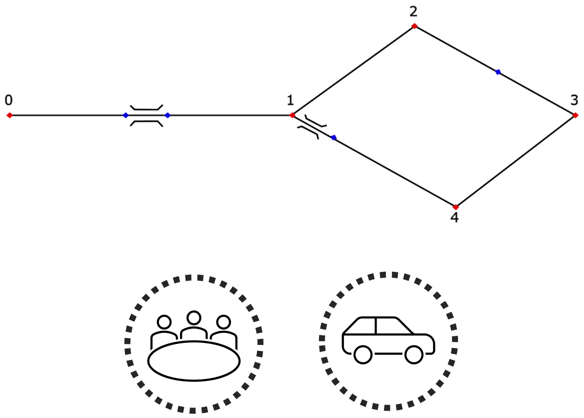
- Based on Deep Q Network
- 2 networks for target and current Q-values
- Reduces overestimation by decoupling

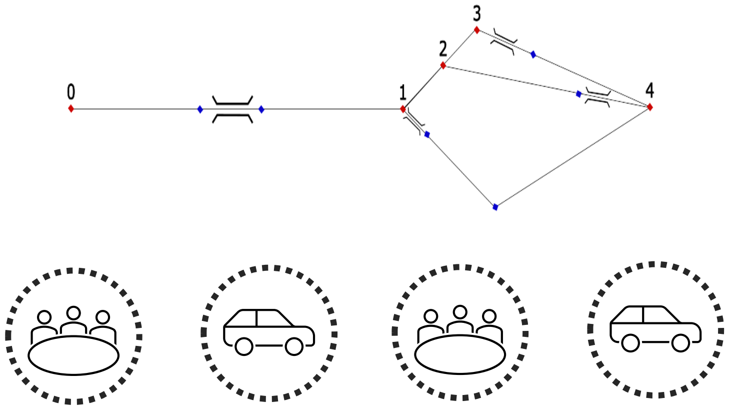
Deep Centralised Multi-agent Actor Critic

- Multi-agent method: Divides the formulation into smaller sub problems
- 2 Networks: Actor and Critic
- Actor network learning based on policy gradient approach.
- Critic network evaluates action by the actor by computing the value function

Scenario Setup







Scenario 1

Multiple individual component setup with one objective i.e., Minimize agency cost.

Scenario 2

Multiple component setup as a network with two objectives i.e., Minimize agency and user cost.

Scenario 3

Entire Transport Network with the goal to minimize multiple objectives for efficient management.

Agenda

01

Introduction

02

Research Overview

03

Environment modelling

04

DRL and scenario setup

05

Results and Discussions

TU Delft

Scenario 1

- Environment description:
  - No. of components: 5 road segments
  - Starting state: Always intact state
  - Objective: Minimise agency cost

- Goal:
  - Interaction between all POMDP components
  - Impact of budget on policy and the costs

Scenario 1

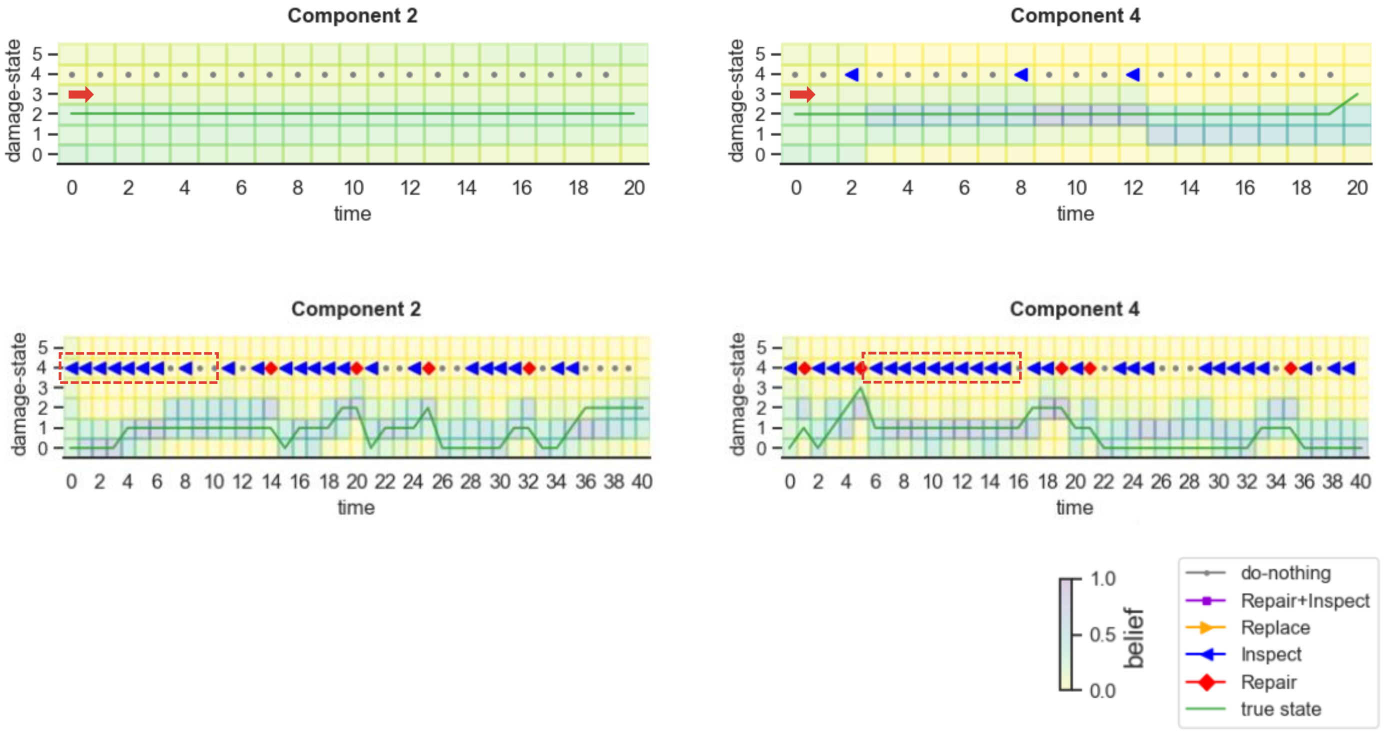
- Algorithm used: DDQN
  - Simplified environment
  - Manageable action and state space

- Parameters
  - Balanced batch size
  - Slightly discounted future rewards
  - Lower target network reset

Parameters	Value
Algorithm	DDQN
Number of episodes	15_000
Memory Capacity	8_000
Batch size	64
Discount factor	0.98
Optimizer	'Adam’
Architecture	[64, 64]
Target network reset	30
Learning rate	0.001
Exploration strategy	Epsilon greedy

Scenario 1 | Run 1

- Goal:
  - Tune the parameters
  - Determine the maximum budget to allocate



- Unaggressive deterioration rates
  - Remains in same state
- Continuous inspection actions

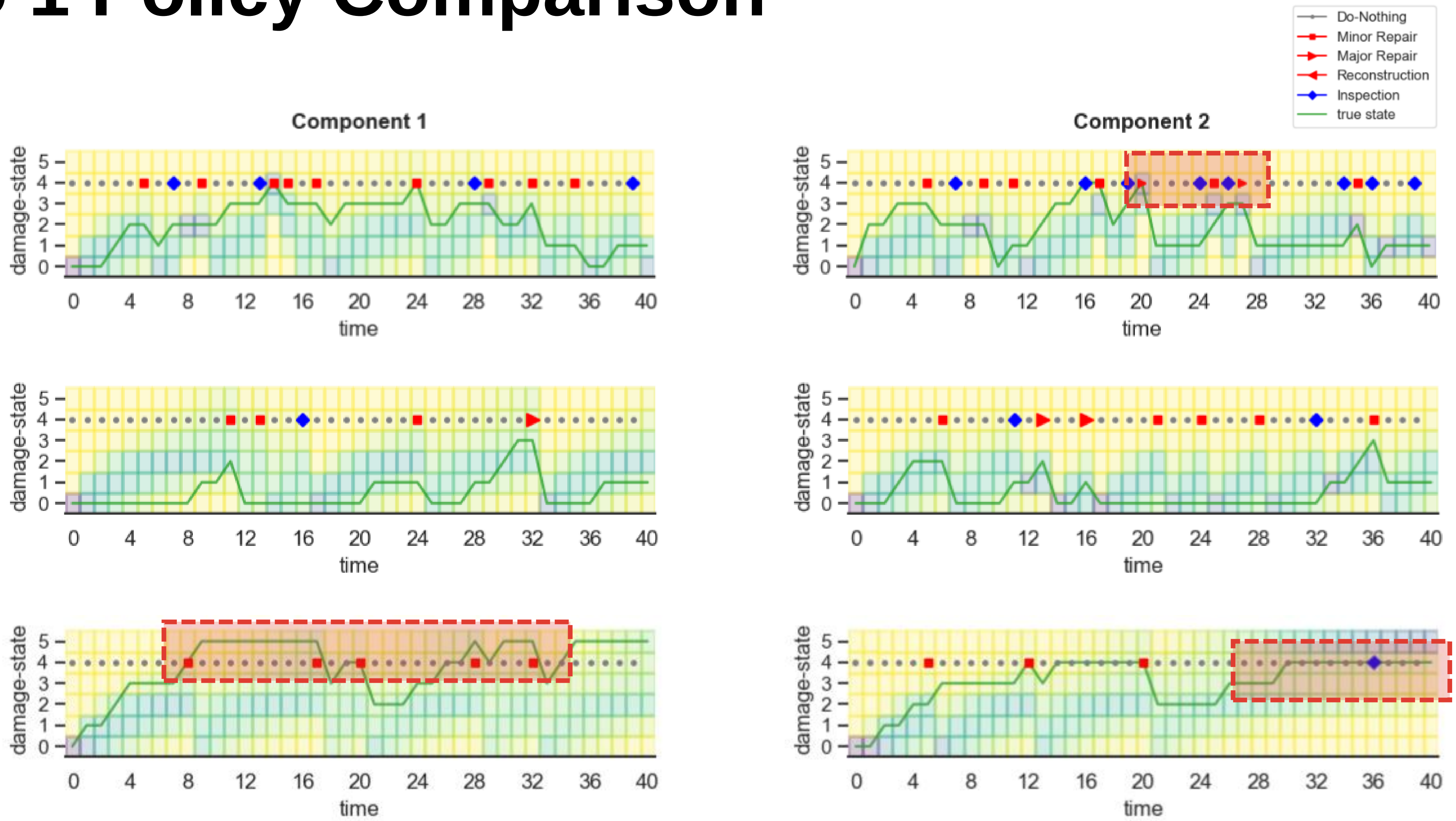
TU Delft

Scenario 1 Policy Comparison

100% Budget

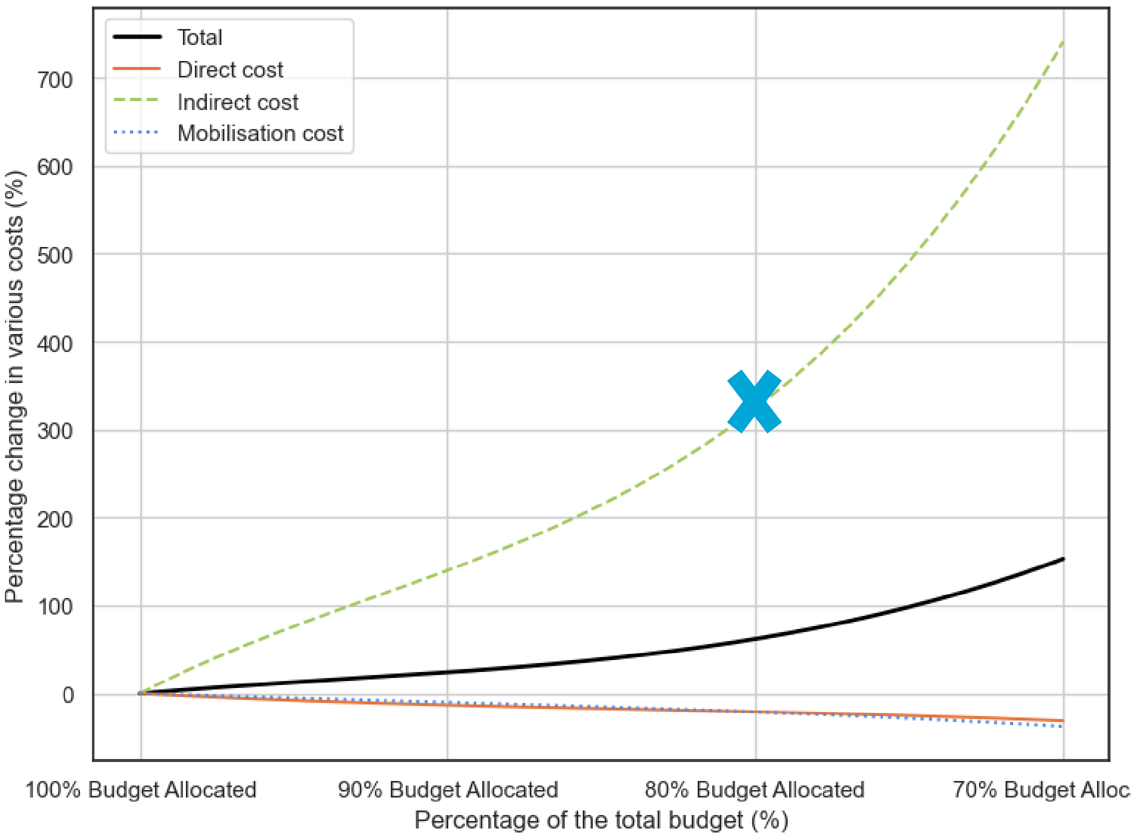
80% Budget

70% Budget



Results

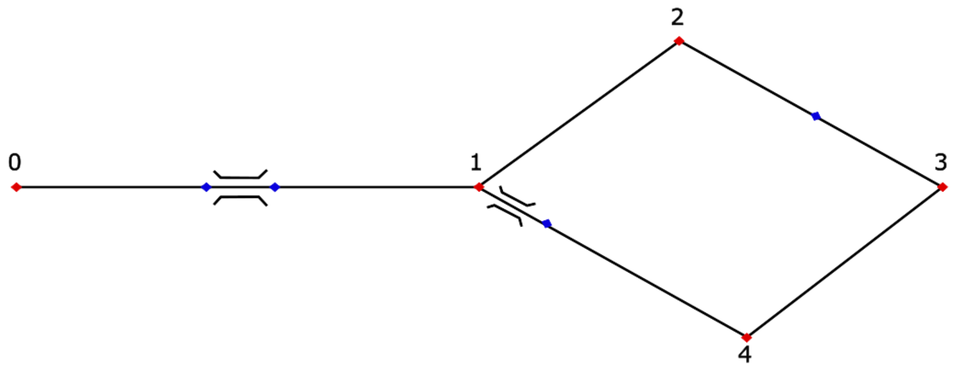
	90%	80%	70%
Total cost	24%	62%	153%
Direct	-13%	-21%	-31%
Mobilising	-10%	-21%	-37%
Indirect	140%	324%	741%



TU Delft

Scenario 2

- Environment description:
  - No. of components: 9 segments (2 bridge, 7 road)
  - Starting state: Non intact state
  - Objective: Minimise agency and user cost



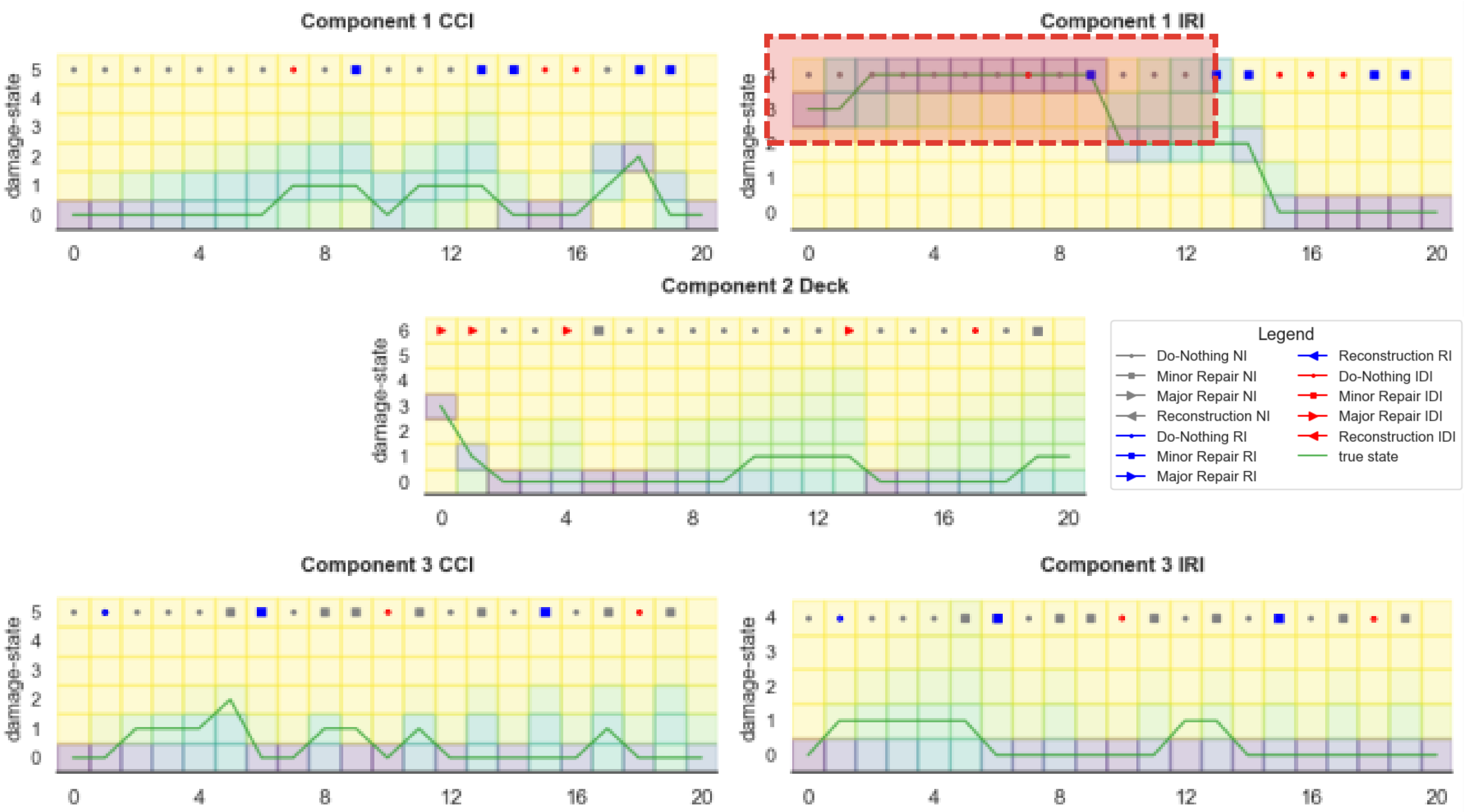
- Goal:
  - Interaction between segments with traffic model
  - Impact of competing objectives

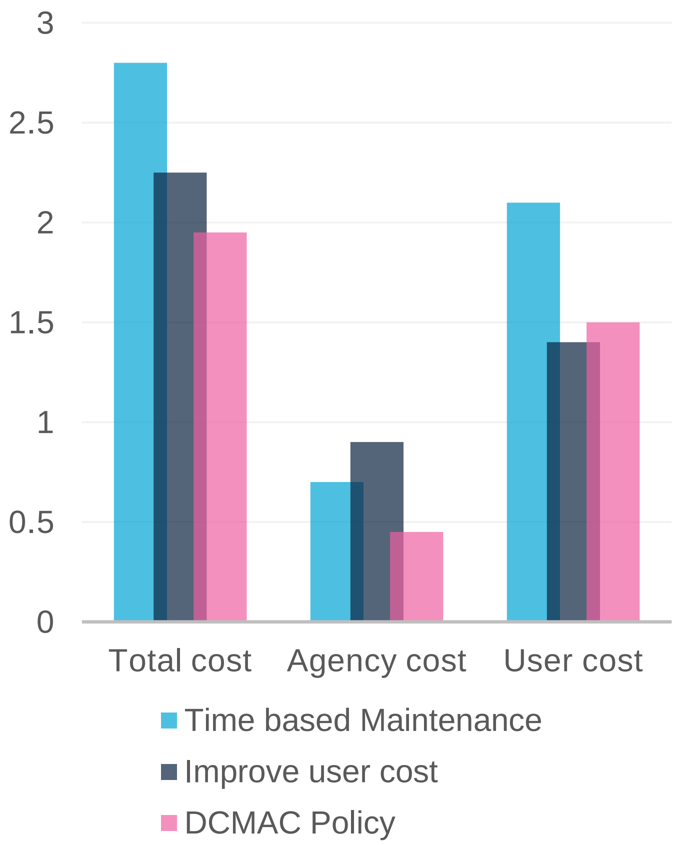
Scenario 2

- Algorithm used: DCMAC
  - Complex environment: large action and state space
  - Scalability
- Parameters
  - Standard batch size and optimizer
  - Lower learning rate
  - Large memory capacity

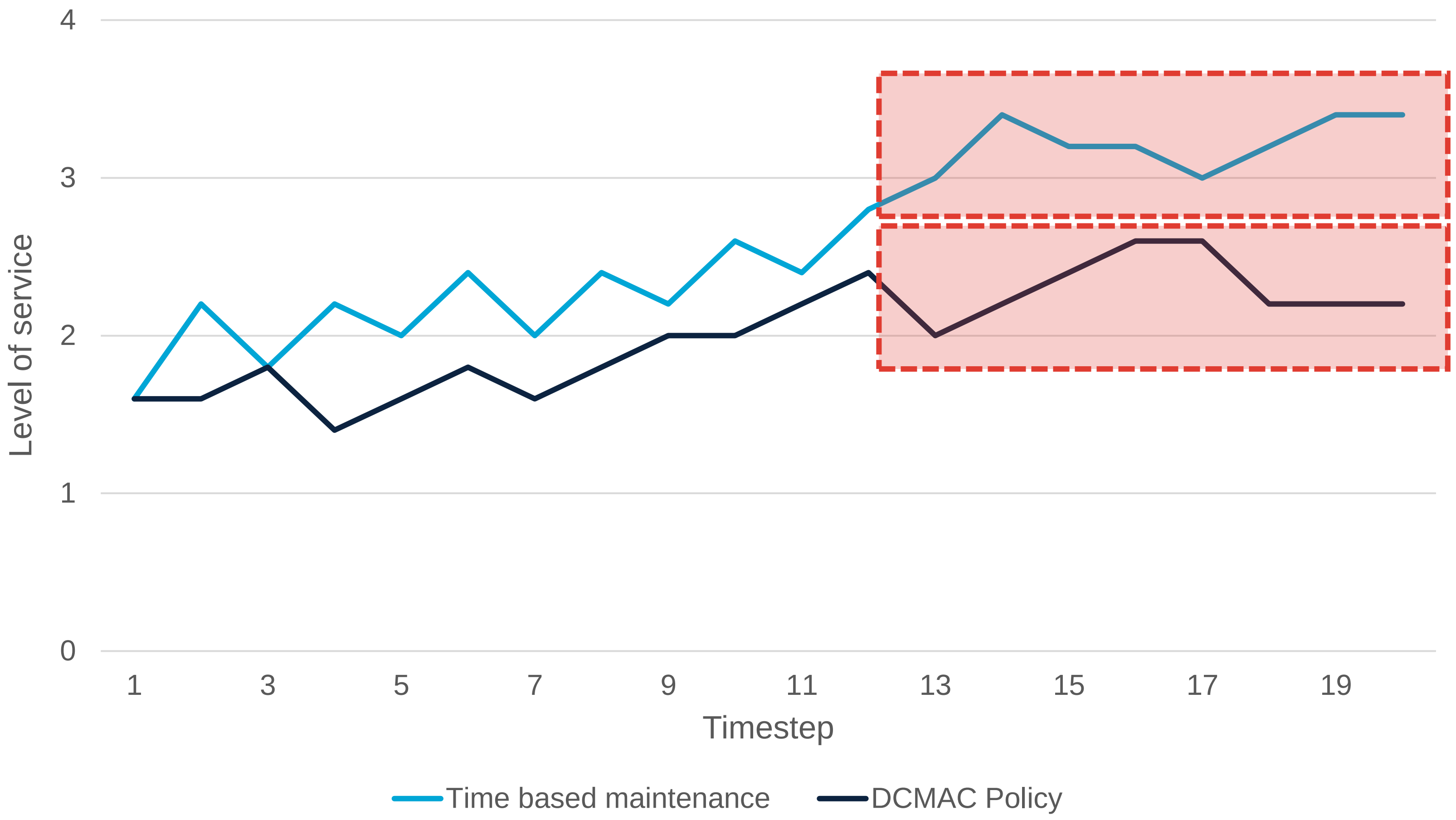
Parameters	Value
Algorithm	DCMAC
Number of episodes	21_000
Batch size	64
Discount factor	0.98
Optimizer	‘Adam’
Actor Architecture	[32, 32]
Actor Learning type	Linear LR
Actor Learning rate	0.0001
Critic Architecture	[64, 64]
Actor Learning rate	0.005
Exploration strategy	Epsilon greedy

Policy Realisation





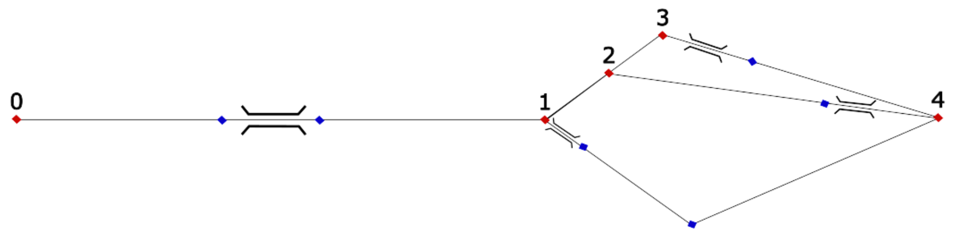
Network congestion



Scenario 3

- Environment description:
  - No. of components: 12 segments (3 bridge, 9 road)
  - Starting state: Non intact state
  - Objective: Minimise all 4 metrics

- Goal:
  - Impact of competing metrics
  - Impact of decision makers preference



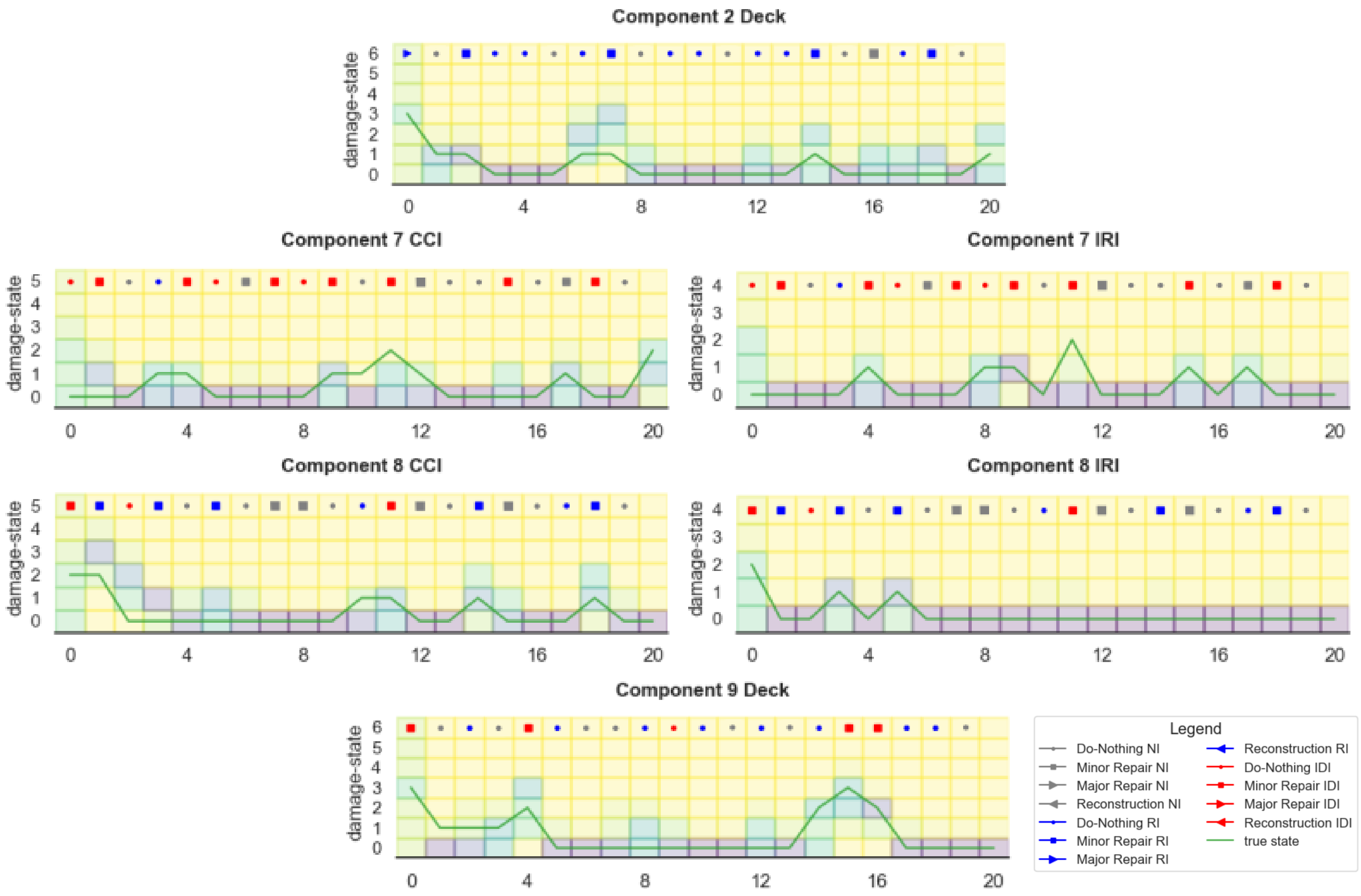
Scenario 3

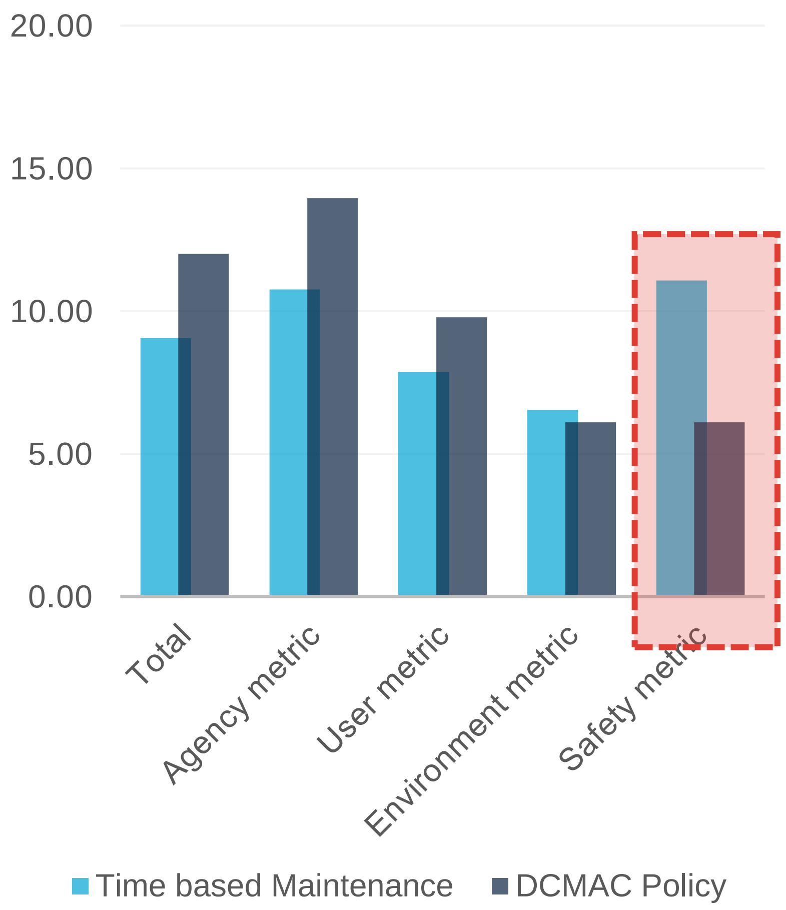
- Algorithm used: DCMAC
  - Complex environment: large action and state space
  - Scalability

- Parameters
  - Increase in number of episodes
  - Large memory capacity

Parameters	Value
Algorithm	DCMAC
Number of episodes	50_000
Batch size	64
Discount factor	0.99
Optimizer	‘Adam’
Actor Architecture	[32, 32]
Actor Learning type	Linear LR
Actor Learning rate	0.0001
Critic Architecture	[64, 64]
Actor Learning rate	0.005
Exploration strategy	Epsilon greedy

Policy Realisation





Reflections (Environment Modelling)

- Segmentation and graph-based Approach
  - Effective policy development by addressing distinct characteristics of each segment
  - Simplifies identification of bottlenecks and critical paths
- Possibility to work at multiple scales; macro, meso-, micro
- Selection of nodes and links critical
- Traffic model enable optimization for free flow, and congestion mitigation

Reflections (RL Experiments)

- State augmentation can cause issues in convergence and policy stability
- Using a Multi attribute utility model improved convergence
  - Simplifying and scaling factors
- ‘MSE’ loss for Critic and ‘Cross Entropy’ loss for actor gave the best results
- DCMAC outperforms DDQN in efficiency and scalability
  - Distributed nature increases robustness

Limitations

- A hypothetical network with data, and values from various regions in the US used.
- Segments deteriorate independently
- Static traffic for the entire year
- Max values for the utility model are assumed
- High computational time constraints
- Excludes important objectives like society and interdependence with other urban systems

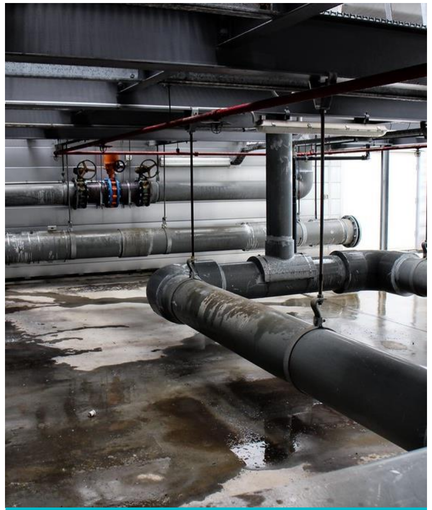
Future Research

- Environment Modelling perspective-
  - Establishing spatial and structural correlations between components
  - Expansion of scope to include decommissioning and EOL scenarios.
  - Identifying critical areas, determining the need for new paths
  - Cross-network management.

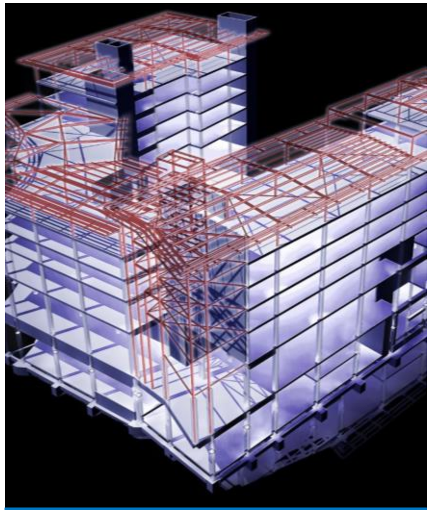
- Computation perspective-
  - Exploring Graph Neural Networks
  - Method to preserve and reuse learned policy with small updates
  - Different loss functions

Prospects of framework in Architecture and the Built Environment

Modular construction and renovation



Building Management Systems



Construction Management

On-site Robot construction

Interdependent urban systems

Thank you for your attention